



PCT

## RAW SEQUENCE LISTING

DATE: 08/26/2004

PATENT APPLICATION: US/10/505,316

TIME: 16:10:16

Input Set : D:\1242-50-3 PCT-US.ST25.txt

Output Set: N:\CRF4\08262004\J505316.raw

3 <110> APPLICANT: Vanderbilt University  
 4 Case Western Reserve University  
 5 Mount, David B  
 6 Romero, Michael  
 8 <120> TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF SLC26A7 and SLC26A9 ANION  
 9 EXCHANGERS  
 11 <130> FILE REFERENCE: 1242/50/3 PCT-US  
 C--> 13 <140> CURRENT APPLICATION NUMBER: US/10/505,316  
 C--> 13 <141> CURRENT FILING DATE: 2004-08-20  
 13 <150> PRIOR APPLICATION NUMBER: US 60/360,287  
 14 <151> PRIOR FILING DATE: 2002-02-28  
 16 <150> PRIOR APPLICATION NUMBER: PCT/US03/06220  
 17 <151> PRIOR FILING DATE: 2003-02-28  
 19 <160> NUMBER OF SEQ ID NOS: 33  
 21 <170> SOFTWARE: PatentIn version 3.2  
 23 <210> SEQ ID NO: 1  
 24 <211> LENGTH: 1971  
 25 <212> TYPE: DNA  
 26 <213> ORGANISM: Homo sapiens  
 28 <400> SEQUENCE: 1  
 29 atgacaggag caaagaggaa aaagaaaagc atgctttgga gcaagatgca tccccccag 60  
 31 tgtgaagaca ttatacagt gtgtagaagg cgactgcca ttttgattg ggcaccacat 120  
 33 tacaatctga aagaaaactt gcttcagac actgtgtctg ggataatgtt ggcagttcaa 180  
 35 caggtgacct aaggattggc ctttgcctgt ctctcatctg tgcacccagt gtttggttta 240  
 37 tatgggtctc tgtttcctgc cataatttat gccatatgtt gaatgggaca tcatgttgcc 300  
 39 acaggcacct ttgacctgac atccttaata tcagccaacg ccgtggaacg gattgtccct 360  
 41 cagaacatgc agaattctcac cacacagagt aacacaagcg tgctgggctt atccgacttt 420  
 43 gaaatgcaaa ggatccacgt tgcctgcagc gtttccttct tgggaggtgt gattcagggtg 480  
 45 gccatgtttg tgctgcaact gggcagtgcc acatttgttg tcacagagcc tgtgatcagc 540  
 47 gcaatgacaa ctggggctgc caccatgtg gtgacttcac aagtcaaata tctcttgga 600  
 49 atgaaaatgc catatatatc cggaccactt ggattctttt atatttatgc atatgttttt 660  
 51 gaaaacatca agtctgtgcg actggaagca ttgcttttat ccttgctgag cattgtggtc 720  
 53 cttgttcttg ttaaagagct gaatgaacag tttaaaagga aaattaaagt tgttcttcct 780  
 55 gtagatntag ttttgattat tgctgcatca tttgcttggt attgcaccaa tatggaaaac 840  
 57 acatatggat tagaagtagt tggatcatatt ccacaaggaa ttccctcacc tagagctccc 900  
 59 ccgatgaaca tcctctctgc ggtgatcact gaagctttcg gagtggcact tgtaggtat 960  
 61 gtggcctcac tggtcttgc tcaaggatct gccaaaaaat tcaaatattc aattgatgac 1020  
 63 aaccaggaaat ttttgcccca tggcctcagc aatatagttt cttcattttt cttctgcata 1080  
 65 ccaagtgtct ctgccatggg aaggacggct ggctgtaca gcacaggagc gaagacacag 1140  
 67 gtggcttgct taatatcttg cattttcgtc cttatagtca tctatgcaat aggaccttg 1200  
 69 ctttactggc tgcccatgtg tgccttgca agcattattg ttgtgggact gaagggaatg 1260  
 71 ctaatacagt tccgagattt aaaaaaatat tggaatgtgg ataaaatcga ttggggaata 1320  
 73 tgggtcagta catatgtatt tacaatatgc tttgctgcca atgtgggact gctggttggt 1380

ENTERED

## RAW SEQUENCE LISTING

DATE: 08/26/2004

PATENT APPLICATION: US/10/505,316

TIME: 16:10:16

Input Set : D:\1242-50-3 PCT-US.ST25.txt

Output Set: N:\CRF4\08262004\J505316.raw

```

75 gttgttttga ccatagctat agtgatagga cgcttcccaa gagcaatgac tgtaagtata 1440
77 aaaaatatga aagaaatgga atttaaagtg aagacagaaa tggacagtga aaccctgcag 1500
79 caggtgaaaa ttatctcaat aaacaacccg cttgttttcc tgaatgcaaa aaaattttat 1560
81 actgatttaa tgaacatgat ccaaaaggaa aatgcctgta atcagccact tgatgatata 1620
83 agcaagtgtg aacaaaacac attgcttaat tccctatcca atggcaactg caatgaagaa 1680
85 gcttcacagt cctgccctaa tgagaagtgt tatttaatcc tggattgcag tggatttacc 1740
87 ttttttgact attctggagt ctccatgctt gttgagggtt acatggactg taaaggcagg 1800
89 agtgtggatg tattgttagc ccattgtaca gcttccttga taaaagcaat gacgtattat 1860
91 ggaaacctag actcagagaa accaattttt tttgaatcgg tatctgctgc aataagtcac 1920
93 atccattcaa ataagaattt gagcaaactc agtgaccaca gtgaagtctg a 1971
96 <210> SEQ ID NO: 2
97 <211> LENGTH: 656
98 <212> TYPE: PRT
99 <213> ORGANISM: Homo sapiens
101 <400> SEQUENCE: 2
103 Met Thr Gly Ala Lys Arg Lys Lys Arg Ser Val Leu Trp Gly Lys Met
104 1 5 10 15
107 His Thr Pro His Arg Glu Asp Ile Lys Gln Trp Cys Lys Arg Arg Leu
108 20 25 30
111 Pro Ile Leu Glu Trp Ala Pro Gln Tyr Asn Leu Lys Glu Asn Leu Leu
112 35 40 45
115 Pro Asp Thr Val Ser Gly Ile Met Leu Ala Val Gln Gln Val Ala Gln
116 50 55 60
119 Gly Leu Ser Phe Ala Met Leu Ser Ser Val His Pro Val Phe Gly Leu
120 65 70 75 80
123 Tyr Gly Ser Leu Phe Pro Ala Ile Ile Tyr Ala Ile Phe Gly Met Gly
124 85 90 95
127 Arg His Val Ala Thr Gly Thr Phe Ala Leu Thr Ser Leu Ile Ser Ala
128 100 105 110
131 Asn Ala Val Glu Arg Leu Val Pro Gln Ser Ser Arg Asn Leu Thr Thr
132 115 120 125
135 Gln Ser Asn Ser Ser Val Leu Gly Leu Ser Glu Phe Glu Leu Gln Arg
136 130 135 140
139 Ile Gly Val Ala Ala Ala Val Ser Phe Leu Gly Gly Val Ile Gln Leu
140 145 150 155 160
143 Val Met Phe Val Leu Gln Leu Gly Ser Ala Thr Phe Leu Leu Thr Glu
144 165 170 175
147 Pro Val Ile Ser Ala Met Thr Thr Gly Ala Ala Thr His Val Val Thr
148 180 185 190
151 Ser Gln Val Lys Tyr Leu Leu Gly Ile Lys Met Pro Tyr Ile Ser Gly
152 195 200 205
155 Pro Leu Gly Phe Phe Tyr Ile Tyr Ala Tyr Val Phe Glu Asn Ile Lys
156 210 215 220
159 Ser Val Gln Leu Glu Ala Leu Leu Leu Ser Leu Leu Ser Ile Ile Val
160 225 230 235 240
163 Leu Val Leu Val Lys Glu Leu Asn Glu Gln Phe Lys Arg Lys Ile Lys
164 245 250 255
167 Val Val Leu Pro Val Asp Leu Val Leu Ile Ile Ala Ala Ser Phe Ala
168 260 265 270

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/505,316

DATE: 08/26/2004

TIME: 16:10:16

Input Set : D:\1242-50-3 PCT-US.ST25.txt

Output Set: N:\CRF4\08262004\J505316.raw

```

171 Cys Tyr Cys Thr Asn Met Glu Asn Thr Tyr Gly Leu Glu Val Val Gly
172      275      280      285
175 His Ile Pro Asn Gly Ile Pro Pro Pro Arg Ala Pro Pro Met Asn Ile
176      290      295      300
179 Leu Ser Ala Val Leu Thr Glu Ala Phe Gly Val Ala Leu Val Gly Tyr
180 305      310      315      320
183 Val Ala Ser Leu Ala Leu Ala Gln Gly Ser Ala Lys Lys Phe Lys Tyr
184      325      330      335
187 Ser Val Asp Asp Asn Gln Glu Phe Leu Ala His Gly Leu Ser Asn Val
188      340      345      350
191 Ile Pro Ser Phe Leu Phe Cys Ile Pro Ser Ala Ala Ala Met Gly Arg
192      355      360      365
195 Thr Ala Gly Leu Tyr Ser Thr Gly Ala Lys Thr Gln Val Ala Cys Leu
196      370      375      380
199 Ile Ser Cys Ile Phe Val Leu Ile Val Ile Tyr Ala Ile Gly Pro Leu
200 385      390      395      400
203 Leu Tyr Trp Leu Pro Met Cys Val Leu Ala Ser Ile Ile Val Val Gly
204      405      410      415
207 Leu Lys Gly Met Leu Ile Gln Phe Arg Asp Leu Lys Lys Tyr Trp Asn
208      420      425      430
211 Val Asp Lys Ile Asp Trp Gly Ile Trp Ile Ser Thr Tyr Ile Phe Thr
212      435      440      445
215 Ile Cys Phe Ala Ala Asn Val Gly Leu Leu Phe Gly Val Ile Cys Thr
216      450      455      460
219 Ile Ala Ile Val Leu Gly Arg Phe Pro Arg Ala Lys Thr Leu Ser Ile
220 465      470      475      480
223 Thr Asp Met Lys Glu Met Glu Leu Lys Val Lys Thr Glu Met His Asp
224      485      490      495
227 Glu Thr Ser Gln Gln Ile Lys Ile Ile Ser Ile Asn Asn Pro Leu Val
228      500      505      510
231 Phe Leu Asn Ala Lys Lys Phe Ser Ala Asp Leu Met Lys Ile Ile Leu
232      515      520      525
235 Lys Glu Ser Asp Ser Asn Gln Pro Leu Asp Asp Val Ser Lys Cys Glu
236      530      535      540
239 Gln Asn Thr Leu Leu Ser Ser Leu Ser Asn Gly Asn Cys Asn Glu Glu
240 545      550      555      560
243 Ala Ser Gln Pro Cys Ser Ser Glu Lys Cys Ser Leu Val Leu Asn Cys
244      565      570      575
247 Ser Gly Leu Thr Phe Phe Asp Tyr Thr Gly Val Ser Thr Leu Val Glu
248      580      585      590
251 Leu Tyr Leu Asp Cys Lys Ser Arg Ser Val Asp Val Phe Leu Ala Asn
252      595      600      605
255 Cys Thr Ala Ser Leu Ile Lys Ala Met Thr Tyr Tyr Gly Asp Leu Asp
256      610      615      620
259 Thr Glu Lys Pro Ile Phe Phe Asp Ser Val Pro Ala Ala Ile Thr Ile
260 625      630      635      640
263 Ile Gln Ser Asn Lys Asn Leu Ser Lys Ala Ser Asp His Ser Glu Val
264      645      650      655
267 <210> SEQ ID NO: 3

```

## RAW SEQUENCE LISTING

DATE: 08/26/2004

PATENT APPLICATION: US/10/505,316

TIME: 16:10:16

Input Set : D:\1242-50-3 PCT-US.ST25.txt

Output Set: N:\CRF4\08262004\J505316.raw

```

268 <211> LENGTH: 2289
269 <212> TYPE: DNA
270 <213> ORGANISM: Mus musculus
272 <400> SEQUENCE: 3
273 tactacagga ggggaatctga aaatctgaaa aaatgacggg agcaaagagg aaaaagagaa      60
275 gcgtgctgtg gggcaagatg cacaccccc accgagaaga cattaagcag tgggtgaaga      120
277 ggcgcctacc tatcttgga tgggcaccgc agtacaatct gaaggaaaac ctgcttccag      180
279 aactgtgtgc tgggataatg ttggcagttc aacagggtggc acaagggctg tcttctgcta      240
281 tgctttcatc tgtacacca gtttttggtt tatatggatc tctgtttcca gccatcattt      300
283 atgctatatt tgggaatgggg cgccatggtt ccacaggcac ctttgccttg acatccttga      360
285 tatcagccaa cgctgtggaa cgactgggtc ctcagagcag caggaacctc accacacaga      420
287 gcaactcaag tgtgttgggc ttatccgagt ttgagctgca gagaatcggg gttgcagcgg      480
289 ctgtttcttt cttgggtgga gtgattcagc tggatcatgt tgtgctgcag ctgggcagtg      540
291 ccacattcct gcttacagag cctgtgatca gcgccatgac cacagggggc gccacccatg      600
293 tctgtacatc acaagtcaag tatctcttgg gaatcaaaat gccatatata tccggaccac      660
295 tgggattctt ttatatattt gcgtatgtct ttgagaacat caagtctgtc cagctggaag      720
297 cactgctctt atctttgctg agcatcattg tccttgttct cgtaaagag ctgaatgaac      780
299 agtttaaaag aaaaattaaa gttgttcttc ctgtcgactt agttttgatc attgccgcct      840
301 catttgcttg ctattgtacc aacatggaaa acacatacgg attagaagta gttgggcaca      900
303 ttccaaacgg gattcctcct ccccggtgtc cccaatgaa catcctctct gcagtgtca      960
305 ctgaagcttt cggagttgca cttgtaggct atgtggcctc gctggctctc gcacaaggat     1020
307 ctgccaagaa gttcaaatat tcagttgatg acaaccagga atttttggcc catggcctca     1080
309 gcaacgtgat ccttctattt ctcttctgca tcccaagcgc tgccgccatg ggaaggacgg     1140
311 ctggtttgta cagcaccggc gcaaagacgc aggtggcttg tctaatactt tgcattttcg     1200
313 tccttattgt catctatgca ataggacctc tgctgtactg gctgcccatg tgcgttcttg     1260
315 caagtattat tgttgtgggc ctgaaggga tgttaataca gttccgggat ttaaaaaaat     1320
317 actggaatgt ggataaaatc gattggggca tatggatcag tacttacatt tttaacaatat     1380
319 gctttgctgc caatgtggga ttgttgtttg gtgttatctg caccatagct atcgtgttag     1440
321 gacgtttccc aaggggcaag actttaagca taacagacat gaaagaaatg gaattgaagg     1500
323 tgaagacaga aatgcatgat gaaacctcgc aacagataaa aatcatctca ataaacaatc     1560
325 cacttgtttt cctgaatgca aagaaattta gtgctgattt gatgaaaata atcctaaagg     1620
327 aaagtgcacg caaccaacca cttgatgatg tcagcaagtg tgaacagaac accttgctca     1680
329 gttccctgtc caacggcaac tgcaatgaag aggcctccca gccctgctcc agcgagaagt     1740
331 gttctttggg cctgaattgc agtggcttga ccttttttga ctacacagga gtctccacgc     1800
333 ttgttgagct ctaccttgat tgcaagagca ggagtgtgga tgtgttctta gccaaactgta     1860
335 ctgcatcctt gataaaagca atgacatact atggagacct agacaccgaa aaaccaattt     1920
337 tttttgactc ggtacctgct gcgataacta tcatccaatc aaataagaat ttgagcaagg     1980
339 ccagtgacca cagtgaagtc tgagacccat gtgttgaaat gcacgcctca tcttttagcag     2040
341 ctgccccaaa gaggccagat cctggcattt tgcacacatt ttggtgtata tacagtgtat     2100
343 cactcagtat gaatgatatg acttggcaac tgcatagcca ctggccaaga cttgctgacc     2160
345 ttctattttt tctaattttc ctttaataaac agattcattt agttatttca tataggcatc     2220
347 aatatgcatt tagctttagt ttactgaaca gtaaacatgc ttttatttct tgcaaaaaaa     2280
349 aaaaaaaaaa
352 <210> SEQ ID NO: 4
353 <211> LENGTH: 656
354 <212> TYPE: PRT
355 <213> ORGANISM: Mus musculus
357 <400> SEQUENCE: 4
359 Met Thr Gly Ala Lys Arg Lys Lys Arg Ser Val Leu Trp Gly Lys Met

```

## RAW SEQUENCE LISTING

DATE: 08/26/2004

PATENT APPLICATION: US/10/505,316

TIME: 16:10:16

Input Set : D:\1242-50-3 PCT-US.ST25.txt

Output Set: N:\CRF4\08262004\J505316.raw

```

360 1           5           10           15
363 His Thr Pro His Arg Glu Asp Ile Lys Gln Trp Cys Lys Arg Arg Leu
364           20           25           30
367 Pro Ile Leu Glu Trp Ala Pro Gln Tyr Asn Leu Lys Glu Asn Leu Leu
368           35           40           45
371 Pro Asp Thr Val Ser Gly Ile Met Leu Ala Val Gln Gln Val Ala Gln
372           50           55           60
375 Gly Leu Ser Phe Ala Met Leu Ser Ser Val His Pro Val Phe Gly Leu
376 65           70           75           80
379 Tyr Gly Ser Leu Phe Pro Ala Ile Ile Tyr Ala Ile Phe Gly Met Gly
380           85           90           95
383 Arg His Val Ala Thr Gly Thr Phe Ala Leu Thr Ser Leu Ile Ser Ala
384           100          105          110
387 Asn Ala Val Glu Arg Leu Val Pro Gln Ser Ser Arg Asn Leu Thr Thr
388           115          120          125
391 Gln Ser Asn Ser Ser Val Leu Gly Leu Ser Glu Phe Glu Leu Gln Arg
392           130          135          140
395 Ile Gly Val Ala Ala Ala Val Ser Phe Leu Gly Gly Val Ile Gln Leu
396 145           150           155           160
399 Val Met Phe Val Leu Gln Leu Gly Ser Ala Thr Phe Leu Leu Thr Glu
400           165          170          175
403 Pro Val Ile Ser Ala Met Thr Thr Gly Ala Ala Thr His Val Val Thr
404           180          185          190
407 Ser Gln Val Lys Tyr Leu Leu Gly Ile Lys Met Pro Tyr Ile Ser Gly
408           195          200          205
411 Pro Leu Gly Phe Phe Tyr Ile Tyr Ala Tyr Val Phe Glu Asn Ile Lys
412           210          215          220
415 Ser Val Gln Leu Glu Ala Leu Leu Leu Ser Leu Leu Ser Ile Ile Val
416 225           230           235           240
419 Leu Val Leu Val Lys Glu Leu Asn Glu Gln Phe Lys Arg Lys Ile Lys
420           245          250          255
423 Val Val Leu Pro Val Asp Leu Val Leu Ile Ile Ala Ala Ser Phe Ala
424           260          265          270
427 Cys Tyr Cys Thr Asn Met Glu Asn Thr Tyr Gly Leu Glu Val Val Gly
428           275          280          285
431 His Ile Pro Asn Gly Ile Pro Pro Pro Arg Ala Pro Pro Met Asn Ile
432           290          295          300
435 Leu Ser Ala Val Leu Thr Glu Ala Phe Gly Val Ala Leu Val Gly Tyr
436 305           310           315           320
439 Val Ala Ser Leu Ala Leu Ala Gln Gly Ser Ala Lys Lys Phe Lys Tyr
440           325          330          335
443 Ser Val Asp Asp Asn Gln Glu Phe Leu Ala His Gly Leu Ser Asn Val
444           340          345          350
447 Ile Pro Ser Phe Leu Phe Cys Ile Pro Ser Ala Ala Ala Met Gly Arg
448           355          360          365
451 Thr Ala Gly Leu Tyr Ser Thr Gly Ala Lys Thr Gln Val Ala Cys Leu
452           370          375          380
455 Ile Ser Cys Ile Phe Val Leu Ile Val Ile Tyr Ala Ile Gly Pro Leu
456 385           390           395           400

```

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/10/505,316

DATE: 08/26/2004

TIME: 16:10:17

Input Set : D:\1242-50-3 PCT-US.ST25.txt

Output Set: N:\CRF4\08262004\J505316.raw

L:13 M:270 C: Current Application Number differs, Replaced Current Application No

L:13 M:271 C: Current Filing Date differs, Replaced Current Filing Date